Phase 2/3 adaptive design utilizing a Bayesian decision analytic approach to dose selection

Zachary Kulkarni Skrivanek, PhD, Brenda Gaydos, PhD

Eli Lilly and Company

April 28, 2009
Acknowledgements

**Eli Lilly**
Medical Director
- James Anderson, MD
Chief Operating Officer
- Paul Hines, MBA
Medical
- Mary Jane Geiger, MD, PhD
Statistics
- Brenda Gaydos, PhD
- Natalie Hall, MS
IT
- Michael A Bell, MS

**Pharmacokinetics**
- Jenny Chien, PhD;
- Michael Heathman, MS

**Regulatory**
- Rose Derbyshire; EU
- Sharon Myers, PhD; US

**Berry Consultants**
Statistics
- Don Berry, PhD
- Scott Berry, PhD
Outline

1. Background and Motivation
   - Critical Path Initiative
   - Clinical Plan

2. Design
   - Dose Selection
   - Rational for Design Selection

3. Operating Characteristics
   - Type I Error
   - Estimation and Bias
   - Homogeneity
   - Dose Selection

4. Implementation
   - DMC and Sponsor Involvement
Outline

1. Background and Motivation
   - Critical Path Initiative
   - Clinical Plan

2. Design
   - Dose Selection
   - Rational for Design Selection

3. Operating Characteristics
   - Type I Error
   - Estimation and Bias
   - Homogeneity
   - Dose Selection

4. Implementation
   - DMC and Sponsor Involvement
Proposed to FDA in April 06

Opportunity

- Type 2 diabetes is identified by FDA as a growing health crisis
- More experience is needed within specific disease states to understand how innovative approaches can be utilized to increase information and reduce escalating development costs.
- Utilize novel adaptive clinical design approach with modeling to more effectively learn about dose response early, reduce uncertainty, and increase probability of technical success

Potential Benefit

To advance appropriate application of novel design and modeling efforts broadly across the industry

Study design developed in consultation with FDA (3 CPI meetings)
Seamless 2/3 Design in Type 2 Diabetes

Develop a Dose-Response Model
- to select doses for 2/3 design
- to assess power of 2/3 design

Published Comparator Data + Subject data

Postprandial glucose

FPG

HbA1c

Safety/Tolerability

Exposure-Response Models

Phase 2/3
- Utilize DIFFERENT data driven models
- Dose Selection
Clinical Plan

5 Pivotal Trials
- Adaptive Seamless 2/3
- 4 Fixed Designs

2 Exploratory Studies prior to Adaptive Seamless 2/3
- Single Dose Study in healthy volunteers
- Bayesian Adaptive Multiple Dose/Proof of Concept Study in patients
Progression of Dose Range

- Phase 1 in HV: 160 fold, 6 doses
- PoC in Patients: 120 fold, 6 doses
- Seamless Design: 12 fold, 7 doses overlapping concentration
- Dose Selection: 2 fold minimum, up to 2 doses not adjacent
Outline

1 Background and Motivation
   - Critical Path Initiative
   - Clinical Plan

2 Design
   - Dose Selection
   - Rational for Design Selection

3 Operating Characteristics
   - Type I Error
   - Estimation and Bias
   - Homogeneity
   - Dose Selection

4 Implementation
   - DMC and Sponsor Involvement
Phase 2/3 Design Distinguishing Features

Two stage adaptive design with dose selection from patients randomized in Stage 1

- Adaptive Randomization utilized in Stage 1
- Operating characteristics are demonstrated through simulation

Dose selection algorithm that incorporates both efficacy & safety

- NOT based on primary analysis of patients randomized in Stage 1
- Few, if any patients will have final endpoint on the primary analysis at time of dose selection
- Timing of dose decision is NOT driven by a fixed sample size
- Decision occurs when sufficient confidence is achieved
- Dose selection must be ENDORSED by the external Data Monitoring Committee (DMC) and then the Lilly Independent Review Committee (IRC)
Objectives of Study

1. Identify up to two doses (Low and High doses) that have a high probability of meeting criteria for safety and efficacy.

2. Demonstrate that these doses show robust glycemic control compared to Active Comparator (AC) and placebo in patients with Type 2 Diabetes Mellitus at 12 months.

3. Primary objective is to demonstrate noninferiority to the AC for the High dose. Five other secondary objectives are included in the primary analysis. The family-wise Type I error is controlled by a modified tree-gatekeeping strategy.¹

Components of the Adaptive Algorithm

1. Longitudinal modeling - project 4 measures out to endpoints (6 months for safety measures and 12 months for efficacy measures)
2. Dose response modeling - for 4 projected endpoints using Normal Dynamic Linear Model
3. Mathematically map 4 endpoints to \((0, \infty)\) using Clinical Utility Index (CUI)
4. Update randomization probabilities, based on CUI
5. Evaluate decision rules
   1. Predictive power criteria for non-inferiority and superiority
   2. Posterior probability criteria for CUI
6. Predictive power calculation for sample size in Stage 2
Two Stage Design

Stage 1 ($n \leq 400$)

- 7 experimental doses ($d_1, ..., d_7$) and 2 comparators, placebo ($pbo$) and Active Comparator ($AC$)
- burn-in period of patients 5 per arm followed by bi-weekly adaptive randomization
- adapting on 3 safety ($S1, S2$ and $S3$) and 1 efficacy ($E1$) endpoints
- after 200 patients enrolled, assess decision alternatives
  - start Stage 2
  - stop for "futility" (efficacy and safety)
  - continue in Stage 1
Two Stage Design

Stage 2 (add approximately 800 patients)
- up to two doses, \textit{pbo} and \textit{AC}
- $\geq 70\%$ of patients in each arm from Stage 2
- fixed allocation to all arms/fixed sample size
- same study schedule as in Stage 1

Final analyses: frequentist
- Includes data from both stages for all of the arms continuing into Stage 2
- Strong control of Type I Error rate prospectively demonstrated via simulations
Two Stage Design

Stage 2 (add approximately 800 patients)
- up to two doses, *pbo* and *AC*
- \( \geq 70\% \) of patients in each arm from Stage 2
- fixed allocation to all arms/fixed sample size
- same study schedule as in Stage 1

Final analyses: frequentist
- Includes data from both stages for all of the arms continuing into Stage 2
- Strong control of Type I Error rate prospectively demonstrated via simulations
Illustration of Design

- **Adaptive Randomization**
  - Placebo 6 months/Active Control
  - Active Control
  - Dose 1
  - Dose 2
  - Dose 3
  - Dose 4
  - Dose 5
  - Dose 6
  - Dose 7
- **Fixed Randomization**
  - Low Dose
  - High Dose

Randomization Stage 1

Decision Point

Randomization Stage 2

Primary Analysis
Criteria for Dose Selection

**High Dose:** Dose estimated to have the *most benefit*

**Low Dose:** Lowest dose estimated to have *meaningful benefit*

**Determining Benefit: Clinical Utility Index (CUI)**

- Multiple endpoints transformed into a single index
  - Safety & efficacy endpoints mapped to an index of patient benefit
- Greater CUI index value reflects greater benefit
- Different functional forms possible to define mapping
  - Additive
  - Multiplicative
Building the CUI

Endpoints
• E1: Efficacy relative to active at 12 months
• S1, S2, S3: Safety relative to placebo at 6 months

CUI multiplicative form selected
• Any endpoint value that surpasses a pre-defined threshold level maps to 0
  • CUI index is 0 value
• Any endpoint value that has a neutral effect on benefit assessment maps to 1
  • CUI index does not increase or decrease in value

Validate with decision makers
• Discuss thresholds for risk and benefit with regulators
• Assess possible drug profiles: Does CUI align with decisions?
• Iterate
Adaptive Randomization Schedule

**Objective:** Randomize more patients to

- Dose demonstrating the most benefit (highest utility)
- Lowest dose that exhibits meaningful benefit (lowest dose that has a utility of at least 0.60)

**Results**

- Increase likelihood of patients receive beneficial treatment
- Improve dose selection by learning more about candidate doses for Stage 2
- Increases sample size to doses in therapeutic range
- Ensure minimum exposure across doses to estimate dose response
Getting the Dose Right

Adequate information across the dose range

- Algorithm does not move away quickly from equal allocation
- At least 18 pts per arm expected
- Large burn-in can be inefficient
Fixed Design Comparison

Robust Phase 2 Trial Design

1. 4 doses and pbo
2. 12 or 26 weeks (2 scenarios)
3. no dropouts
4. 50 patients/arm

Assess both fixed design and adaptive design using the same pharmacodynamic models for $E_1, S_1, S_2$ and $S_3$

1. fixed design: 8% or 18% (12 or 26 weeks) successfully selected dose(s)
2. adaptive design: 89% success rate (mean duration 56 weeks, mean total sample size 279 patients)
Fixed Design Comparison

Robust Phase 2 Trial Design

1. 4 doses and pbo
2. 12 or 26 weeks (2 scenarios)
3. no dropouts
4. 50 patients/arm

Assess both fixed design and adaptive design using the same pharmacodynamic models for $E_1$, $S_1$, $S_2$ and $S_3$

1. fixed design: 8% or 18% (12 or 26 weeks) successfully selected dose(s)
2. adaptive design: 89% success rate (mean duration 56 weeks, mean total sample size 279 patients)
Difficult Experimental Situation

High variability in safety measures
Need to study more doses within the narrowed dose range to better characterize benefit risk

It is NOT about Efficacy
- Volume of information and experience in Type 2 diabetes
- Well established mechanism of action
- Availability of validated biomarker for anti-diabetic agents (HbA1c)
Difficult Experimental Situation

Advantage of adaptive randomization based on CUI

Significant improvement demonstrated if more doses studied over a longer duration at a level unprecedented in exploratory development
Totality of Evidence vs. Separate Trials

**Benefits of Seamless Design**
- Long term safety data available sooner
- Increases totality of evidence through continued exposure of Stage 1 patients
- Independent DMC oversight
- More evidence available prior to start of other Phase 3 studies
- Moves non-data generating activities to parallel tasks

**Drawbacks of Seamless Design**
- Restricted access to Stage 1 data
  - ERBs
  - Regulatory agencies
  - Sponsor
- Potential for introduction of operational bias

---

**Adaptive, Seamless Clinical Approach**

- Stage 1
- Stage 2
- Add 1 Phase 3 studies

**Adaptive Design Followed by Phase 3**

- Adaptive Ph 2
- Additional Activities
- Time to enroll additional patients
- 9-12 Months
- Add 1 Phase 3 studies
**Mitigation Steps**

- Data from additional 4 studies to optimize design of Phase 3 protocols
- At least 4 fixed design Phase 3 trials to confirm dose selection
- Operational bias mitigated through design, firewalls, well documented processes

**Conclusions**

- Adaptive design maximizes the use of ongoing information to make decisions while reducing the probability of exposing patients to an ineffective drug or the wrong dose in Phase 3
- The totality of evidence at the decision point, at the start of the Phase 3 trials and at submission is preferable to the separate trial scenario
  - Data from longer exposure to doses are available, enabling better dose decision
  - Data from longer exposure in Stage 1 contribute to superior data package
Outline

1. Background and Motivation
   - Critical Path Initiative
   - Clinical Plan

2. Design
   - Dose Selection
   - Rational for Design Selection

3. Operating Characteristics
   - Type I Error
   - Estimation and Bias
   - Homogeneity
   - Dose Selection

4. Implementation
   - DMC and Sponsor Involvement
Extensive Simulations

**Response Models**
- PK/PD (alternative)
- Empirical statistical models (null or alternative)

**Data Analysis Model**
- ANCOVA
- LOCF
- Tree-gatekeeping
- Nominal alpha

**Trial Performance Metrics**
- Power (Scenario in alternative)
- Type I error (Scenario in null)
- Probability of selecting correct dose(s)
- Probability of stopping for safety
- Probability of stopping for lack of efficacy

**Virtual Patients**
- Study Information
  - Inclusion/exclusion criteria
  - Visit schedule

**Trial Simulator**
- Trial Execution Models
  - Dropout models
  - Accrual models

**Study Design**
- Predictive models:
  - Longitudinal model to predict patient outcomes for ongoing patients
  - Dose response model to assess population effect
- Decision models:
  - CUI

**Operating Characteristics**
- Virtual Results
- Output

**Simulation Report**

Skrivanek, Gaydos

FDA/DIA Statistics Forum
Missingness Scenarios

**Missing Completely At Random (MCAR)**
- varied dropout rate
- constant hazards model
- varied accrual rate
- assumed no safety issues
- conservative null models

**Missing at Random (MAR) Missing Not at Random (MNAR)**
- different dropout rates between arms
- differential hazards models
- covariate dependent dropouts
Type I Error Simulations

- span entire null space
- simulate same scenarios for comparable fixed design
- control for confounding factors to isolate cause of any inflation of Type I error.
- assumed no safety issues
- conservative null models
Simulations demonstrated the Type I error rate was well controlled

<table>
<thead>
<tr>
<th>No Dropouts</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Design &amp; analysis is <em>very conservative</em></td>
</tr>
<tr>
<td>- All estimates were at least 2 standard errors below the targeted 0.025</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dropouts</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Dropouts induced bias due to LOCF</td>
</tr>
<tr>
<td>- Adaptive design as well as or better than fixed design</td>
</tr>
<tr>
<td>- Conservative adjustment in adaptive design help to mitigate bias due to LOCF</td>
</tr>
</tbody>
</table>
Little to No Inflation of Type I Error Rate

- No more than 30% of patients come from Stage 1
- Adaptations based on 3 safety measures 1 efficacy measure
- Forced futility
Estimate of Effect

Bias induced by the adaptive design is dependent on

- True unknown dose response
- Drop out and enrollment rate

Most likely pharmacodynamic response

- Simulations (adjusted for LOCF bias) demonstrated no to small, clinically irrelevant bias
- MAXIMUM of within group and between group across ALL drop-out rates and enrolment rates is $\sim 5\%$ of non-inferiority margin
- Non-inferiority margin for this study small ($< 0.3\%$)
- 95\% confidence interval coverage was maintained
- adjusted 2-sided alpha of 0.04 used to calculated 95\% CI
Mitigating Heterogeneity Across Stages

By Design

- Double-blind study throughout
- Same inclusion/exclusion criteria across stages
- Similar countries, demographically, across stages
- Visit schedule the same across stages
- Firewalls & well documented procedures for interactions between IDMC & ISC

To ensure interpretability

- Fixed randomization in Stage 2
- At least 70% of the patients in each arm of the primary analysis will come from Stage 2
- 4 other Phase 3 trials
- Stage 2 powered to stand on its own
Example Results: narrow therapeutic window

- probability increases with CUI
- sample size increases with CUI
Outline

1. Background and Motivation
   - Critical Path Initiative
   - Clinical Plan

2. Design
   - Dose Selection
   - Rational for Design Selection

3. Operating Characteristics
   - Type I Error
   - Estimation and Bias
   - Homogeneity
   - Dose Selection

4. Implementation
   - DMC and Sponsor Involvement
**Data Flow**

- **sites**
  - eCRF Data
  - Central Lab Data
  - Other Data* (e.g., PK, adjudicated AE data)
  - Pharmacovigilence Data

- **BLINDED**
  - Lilly Sr. Mgt. Designee
  - Recommendations

- **UNBLINDED**
  - Lilly IRC
  - DMC
  - ISC
  - IVRS
  - Safety & Efficacy Data
  - bi-weekly randomization updates

---

*IRC = Internal Review Committee
DMG = Data Movement Group
ISC = Independent Statistical Center
IVRS = Interactive Voice Response System
### Lilly Internal Review Committee (IRC)

**Membership (independent of the clinical development team)**
- 3 Clinicians, 1 statistician

**IRC/DMC Communication**
- Through Senior Management Designee

**Information Review**
- Senior Management Designee may request an internal review of data based on DMC recommendation
- The IRC cannot ask for additional information, but can review the same reports as the DMC
- At the time of Decision point, the IRC will receive the DMC recommendation and review the reports

**Actions of IRC limited to:**
- Agreement with DMC recommendation
- AND/OR Stop the study
Independent Data Monitoring Committee (DMC)

**Composition**
External experts in medicine and biostatistics

**Role:** To review safety data and interim results to make recommendations to the Sponsor designee regarding:

- The performance of the adaptive algorithm during Stage 1
- To discontinue any dose arm(s) due to safety concerns, or stop the study
- To endorse up to 2 doses identified by the algorithm to carry forward into Stage 2 and other Phase 3 trials
- To terminate the study due to futility

The DMC can only recommend doses (or stop the study for futility) based upon the output of the pre-specified algorithm.
Independent Data Monitoring Committee (DMC)

Composition

External experts in medicine and biostatistics

Role: To review safety data and interim results to make recommendations to the Sponsor designee regarding:

- The performance of the adaptive algorithm during Stage 1
- To discontinue any dose arm(s) due to safety concerns, or stop the study
- To endorse up to 2 doses identified by the algorithm to carry forward into Stage 2 and other Phase 3 trials
- To terminate the study due to futility

The DMC can only recommend doses (or stop the study for futility) based upon the output of the pre-specified algorithm.
Summary and Conclusions

- Benefits
- Mitigating Risks
Benefits

- **Rigorous**
  - Planning & Simulations
  - Communication with regulators (iterative planning & simulations)
  - Documentation

- **Improvement in patient assignments**
  - More likely to be assigned to beneficial doses
  - Increased monitoring of patient safety

- **Improvement in dose selection (safety driven)**
  - Without requiring an unprecedented large, long duration exploratory study
  - Does not cut corners
Mitigating Risks

- Adaptations are limited & completely pre-specified
- Conservative Type I error control
- Data integrity maintained
  - Blind maintained / Firewalls
  - DMC construct & Limited sponsor involvement
- Interpretability
  - One of 5 Phase 3 studies
  - Sufficient other available data to optimize design of Phase 3 protocols
  - In seamless design, over 70% of patients randomized in stage 2

Opportunity to gain needed experience

Effort to advance appropriate application of novel design & modeling efforts to improve drug development